RAW SEQUENCE LISTING

- The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10573,600A
Source: 1700
Date Processed by STIC: 8/23/06

ENTERED



IFWO

RAW SEQUENCE LISTING DATE: 08/23/2006
PATENT APPLICATION: US/10/573.600A TIME: 10:23:13

```
1 <110> APPLICANT: The Trustees of the University of Pennsylvania
        Wilson, James M.
 3
        Gao, Guangping
 4
        Alvira, Mauricio R.
        Vandenberghe, Luk H.
 6 <120> TITLE OF INVENTION: Adeno-Associated Virus (AAV) Clades, Sequences, Vectors
 7
         Containing Same, and Uses Therefor
 8 <130> FILE REFERENCE: UPN-P3230PCT
 9 <140> CURRENT APPLICATION NUMBER: US/10/573,600A
10 <141> CURRENT FILING DATE: 2006-03-24
11 <150> PRIOR APPLICATION NUMBER: US 60/508,226
12 <151> PRIOR FILING DATE: 2003-09-30
13 <150> PRIOR APPLICATION NUMBER: US 60/566,546
14 <151> PRIOR FILING DATE: 2004-04-29
15 <160> NUMBER OF SEQ ID NOS: 236
16 <170> SOFTWARE: PatentIn version 3.3
18 <210> SEO ID NO: 1
19 <211> LENGTH: 2211
20 <212> TYPE: DNA
21 <213> ORGANISM: Unknown
22 <220> FEATURE:
23 <223> OTHER INFORMATION: adeno-associated virus, clone hu.31
24 <400> SEQUENCE: 1
25
         atggctgccg atggttatct tccagattgg ctcgaggaca accttagtga aggaattcgc
                                                                                60
         gagtggtggg ctttgaaacc tggagcccct caacccaagg caaatcaaca acatcaagac
26
                                                                               120
27
         aacgctcgag gtcttgtgct tccgggttac aaataccttg gacccggcaa cggactcgac
                                                                               180
28
                                                                               240
         aagggggage eggteaacge ageagaegeg geggeeeteg ageacgaeaa ggeetaegae
29
                                                                               300
         cagcagetea aggeeggaga caaccegtae etcaagtaca accaegeega egeegagtte
30
         caggagcggc tcaaagaaga tacgtctttt gggggcaacc tcgggcgagc agtcttccag
                                                                               360
31
        gccaaaaaga ggcttcttga acctcttggt ctggttgagg aagcggctaa gacggctcct
                                                                               420
32
        ggaaagaaga ggcctgtaga gcagtctcct caggaaccgg actcctccgc gggtattggc
                                                                               480
                                                                               540
33
         aaatcqggtg cacagcccqc taaaaagaga ctcaatttcg gtcagactgg cgacacagag
         tragtrocag accetraace aateggagaa ceteergeag ceeeeteagg tgtgggatet
                                                                               600
34
35
                                                                               660
        cttacaatgg cttcaggtgg tggcgcacca gtggcagaca ataacgaagg tgccgatgga
                                                                               720
36
        gtgggtagtt cctcgggaaa ttggcattgc gattcccaat ggctggggga cagagtcatc
37
        accaccagca cccgaacctg ggccctgccc acctacaaca atcacctcta caagcaaatc
                                                                               780
38
        tccaacagca catctggagg atcttcaaat gacaacgcct acttcggcta cagcacccc
                                                                               840
        tgggggtatt ttgacttcaa cagattccac tgccacttct caccacgtga ctggcagcga
39
                                                                               900
40
        ctcatcaaca acaactgggg attccggcct aagcgactca acttcaagct cttcaacatt
                                                                               960
                                                                              1020
41
        caggicaaag aggitacgga caacaatgga gicaagacca tcgccaataa ccttaccagc
42
                                                                              1080
        acggtccagg tcttcacgga ctcagactat cagctcccgt acgtgctcgg gtcggctcac
43
        gagggetgee teeegeegtt eeeageggae gtttteatga tteeteagta egggtatetg
                                                                              1140
44
        acgettaatg atggaageca ggeegtgggt egttegteet tttactgeet ggaatattte
                                                                              1200
```

RAW SEQUENCE LISTING DATE: 08/23/2006
PATENT APPLICATION: US/10/573,600A TIME: 10:23:13

```
45
                                                                               1260
         ccgtcgcaaa tgctaagaac gggtaacaac ttccagttca gctacgagtt tgagaacgta
46
                                                                               1320
         cetttecata geagetaege teacageeaa ageetggaee gaetaatgaa tecaeteate
47
         gaccaatact tgtactatct ctcaaagact attaacggtt ctggacagaa tcaacaaacg
                                                                               1380
48
         ctaaaattca gtgtggccgg acccagcaac atggctgtcc agggaagaaa ctacatacct
                                                                               1440
49
         ggacccagct accgacaaca acgtgtctca accactgtga ctcaaaacaa caacagcgaa
                                                                               1500
50
         tttgcttggc ctggagcttc ttcttgggct ctcaatggac gtaatagctt gatgaatcct
                                                                               1560
51
         ggacctgcta tggccagcca caaagaagga gaggaccgtt tctttccttt gtctggatct
                                                                               1620
52
         ttaatttttg gcaaacaagg aactggaaga gacaacgtgg atgcggacaa agtcatgata
                                                                               1680
53
                                                                               1740
         accaacgaag aagaaattaa aactactaac ccggtagcaa cggagtccta tggacaagtg
         gccacaaacc accagagtgc ccaagcacag gcgcagaccg gctgggttca aaaccaagga
                                                                               1800
54
55
         atacttccgg gtatggtttg gcaggacaga gatgtgtacc tgcaaggacc catttgggcc
                                                                               1860
56
         aaaattcctc acacggacgg caactttcac ccttctccgc tgatgggagg gtttggaatg
                                                                               1920
57
         aagcacccgc ctcctcagat cctcatcaaa aacacacctg tacctgcgga tcctccaacg
                                                                               1980
58
         gccttcaaca aggacaagct gaactctttc atcacccagt attctactgg ccaagtcagc
                                                                               2040
59
         gtggagatcg agtgggagct gcagaaggaa aacagcaagc gctggaaccc ggagatccag
                                                                               2100
60
         tacacttcca actattacaa gtctaataat gttgaatttg ctgttaatac tgaaggtgta
                                                                               2160
         tatagtgaac cccgccccat tggcaccaga tacctgactc gtaatctgta a
                                                                               2211
63 <210> SEQ ID NO: 2
64 <211> LENGTH: 2211
65 <212> TYPE: DNA
66 <213> ORGANISM: Unknown
67 <220> FEATURE:
68 <223> OTHER INFORMATION: new AAV serotype, clone hu.32
69 <400> SEQUENCE: 2
70
         atggctgccg atggttatct tccagattgg ctcgaggaca ctctctctga aggaataaga
                                                                                 60
71
         cagtggtgga agctcaaacc tggcccacca ccaccaaagc ccgcagagcg gcataaggac
                                                                                120
72
                                                                                180
         gacagcaggg gtcttgtgct tcctgggtac aagtacctcg gacccggcaa cggactcgac
73
                                                                                240
         aagggggage eggteaaege ageagaegeg geggeeeteg ageaegaeaa ggeetaegae
74
                                                                                300
         cagcagetea aggeeggaga caaccegtae etcaagtaca accaegeega egeegagtte
75
         caggagegge teaaagaaga taegtetttt gggggeaace tegggegage agtetteeag
                                                                                360
76
                                                                                420
         gccaaaaaga ggcttcttga acctcttggt ctggttgagg aagcggctaa gacggctcct
77
         ggaaagaaga ggcctgtaga gcagtctcct caggaaccgg actcctccgc gggtattggc
                                                                                480
78
         aaatcgggtt cacagcccgc taaaaagaaa ctcaatttcg gtcagactgg cgacacagag
                                                                                540
79
         teagteeceg acceteaace aateggagaa ceteeegeag eeceeteagg tgtgggatet
                                                                                600
80
         cttacaatgg cttcaggtgg tggcgcacca gtggcagaca ataacgaagg tgccgatgga
                                                                                660
81
         gtgggtagtt cctcgggaaa ttggcattgc gattcccaat ggctggggga cagagtcatc
                                                                                720
82
         accaccagca cccgaacctg ggccctgccc acctacaaca atcacctcta caagcaaatc
                                                                                780
83
         tccaacagca catctggagg atcttcaaat gacaacgcct acttcggcta cagcacccc
                                                                                840
                                                                                900
84
         tgggggtatt ttgacttcaa cagattccac tgccacttct caccacgtga ctggcagcga
85
         ctcatcaaca acaactgggg attccggcct aagcgactca acttcaagct cttcaacatt
                                                                                960
86
         caggtcaaag aggttacgga caacaatgga gtcaagacca tcgccaataa ccttaccagc
                                                                               1020
87
         acggtccagg tcttcacgga ctcagactat cagctcccgt acgtgctcgg gtcggctcac
                                                                               1080
                                                                               1140
88
         gagggetgee teeegeegtt eecageggae gtttteatga tteeteagta egggtatetg
89
         acgcttaatg atgggagcca ggccgtgggt cgttcgtcct tttactgcct ggaatatttc
                                                                               1200
90
         ccgtcgcaaa tgctaagaac gggtaacaac ttccagttca gctacgagtt tgagaacgta
                                                                               1260
                                                                               1320
91
         cetttecata geagetacge teacageeaa ageetggace gaetaatgaa tecaeteate
92
         qaccaatact tgtactatct ctcaaaqact attaacqqtt ctqqacaqaa tcaacaaacq
                                                                               1380
93
         ctaaaattca gcgtggccgg acccagcaac atggctgtcc agggaagaaa ctacatacct
                                                                               1440
                                                                               1500
94
         ggacccagct accgacaaca acgtgtctca accactgtga ctcaaaacaa caacagcgaa
```

RAW SEQUENCE LISTING DATE: 08/23/2006 PATENT APPLICATION: US/10/573,600A TIME: 10:23:13

95	tttgcttggc ctggagcttc				_	1560
96	ggacctgcta tggccagcca					1620
97	ttaatttttg gcaaacaagg					1680
98	accaacgaag aagaaattaa					1740
99	gccacaaacc accagagtgc					1800
100	atacttccgg gtatggtttg					1860
101	aaaattcctc acacggacgg					1920
102	aagcacccgc ctcctcagat					1980
103	gctttcaata aggacaagct					2040
104	gtggagattg agtgggagct					2100
105	tacacttcca actattacaa					2160
106	tatagtgaac cccgccccat	tggcaccaga	tacctgactc	gtaatctgta	a	2211
108 <	210> SEQ ID NO: 3					
109 <	211> LENGTH: 2211					
110 <	212> TYPE: DNA					
111 <	213> ORGANISM: Unknown					
	220> FEATURE:					
	223> OTHER INFORMATION: ad	eno-associat	ted virus, 1	human clone	9	
114 <	400> SEQUENCE: 3					
115	atggctgccg atggttatct	tccagattgg	ctcgaggaca	accttagtga	aggaattcgc	60
116	gagtggtggg ctttgaaacc	tggagcccct	caacccaagg	caaatcaaca	acatcaagac	120
117	aacgctcgag gtcttgtgct	tccgggttac	aaataccttg	gacccggcaa	cggactcgac	180
118	aagggggagc cggtcaacgc	agcagacgcg	gcggccctcg	agcacgacaa	ggcctacgac	240
119	cagcagctca aggccggaga	caacccgtac	ctcaagtaca	accacgccga	cgccgagttc	300
120	caggagcggc tcaaagaaga	tacgtctttt	gggggcaacc	tcgggcgagc	agtcttccag	360
121	gccaaaaaga ggcttcttga	acctcttggt	ctggttgagg	aagcggctaa	gacggctcct	420
122	ggaaagaaga ggcctgtaga	gcagtctcct	caggaaccgg	actcctccgc	gggtattggc	480
123	aaatcgggtg cacagcccgc	taaaaagaga	ctcaatttcg	gtcagactgg	cgacacagag	540
124	tcagtcccag accctcaacc	aatcggagaa	cctcccgcag	cccctcagg	tgtgggatct	600
125	cttacaatgg cttcaggtgg	tggcgcacca	gtggcagaca	ataacgaagg	tgccgatgga	660
126	gtgggtagtt cctcgggaaa	ttggcattgc	gattcccaat	ggctggggga	cagagtcatc	720
127	accaccagca cccgaacctg	ggccctgccc	acctacaaca	atcacctcta	caagcaaatc	780
128	tccaacagcà catctggagg	atcttcaaat	gacáacgcct	acttcggcta	cagcaccccc	840
129	tgggggtatt ttgacttcaa					900
130	ctcatcaaca acaactgggg	attccggcct	aagcgactca	acttcaagct	cttcaacatt	960
131	caggtcaaag aggttacgga	caacaatgga	gtcaagacca	tcgccaataa	ccttaccagc	1020
132	acggtccagg tcttcacgga	ctcagactat	cagctcccgt	acgtgctcgg	gtcggctcac	1080
133	gagggctgcc tcccgccgtt	cccagcggac	gttttcatga	ttcctcagta	cgggtatctg	1140
134	acgcttaatg atggaagcca	ggccgtgggt	cgttcgtcct	tttactgcct	ggaatatttc	1200
135	ccgtcgcaaa tgctaagaac	gggtaacaac	ttccagttca	gctacgagtt	tgagaacgta	1260
136	cctttccata gcagctacgc	tcacagccaa	agcctggacc	gactaatgaa	tccactcatc	1320
137	gaccaatact tgtactatct	ctcaaagact	attaacggtt	ctggacagaa	tcaacaaacg	1380
138	ctaaaattca gtgtggccgg					1440
139	ggacccagct accgacaaca					1500
140	tttgcttggc ctggagcttc					1560
141	ggacctgcta tggccagcca					1620
142	ttaatttttg gcaaacaagg					1680
143	accaacgaag aagaaattaa					1740
144	gccacaaacc accagagtgc					1800
		-	-	_		

RAW SEQUENCE LISTING DATE: 08/23/2006 PATENT APPLICATION: US/10/573,600A TIME: 10:23:13

	aaaattcctc aagcacccgc gccttcaaca gtggagatcg tacacttcca	acacggacgg ctcctcagat aggacaagct agtgggagct actattacaa cccgccccat 4	caactttcac cctcatcaaa gaactctttc gcagaaggaa gtctaataat	gatgtgtacc ccttctccgc aacacacctg atcacccagt aacagcaagc gttgaatttg tacctgactc	tgatgggagg tacctgcgga attctactgg gctggaaccc ctgttaatac	gtttggaatg tcctccaacg ccaagtcagc ggagatccag tgaaggtgta	1860 1920 1980 2040 2100 2160 2211		
	TYPE: DNA								
		ORGANISM: Unknown							
157 <220>									
	OTHER INFOR		w AAV seroty	ype, clone b	ıu.17				
	SEQUENCE: 4								
160				ctcgaggaca			60		
161				aaacccaaag			120		
162				aagtacctcg			180		
163 164				gcggccctcg			240		
				ctgcggtata			300		
165 166				gggggcaacc			360 420		
167				ctggttgagg cagcgttctc			420		
168					-		540		
169				agactcaact gaaccccccg			600		
170				ccaatggcag			660		
171	ggagtgggta						720		
172				cccacctaca			780		
173				aacgacaaca		_	840		
174				cactgccact			900		
175				cccaagagac			960		
176				ggcaccaaga			1020		
177				taccagetee			1080		
178	caccagggct						1140		
179	ctgactctga						1200		
180	tttccttctc						1260		
181	gtgccttttc						1320		
182	atcgaccagt	acctgtacta	cctgtctcgg	actcagtcca	cgggaggtac	cgcaggaact	1380		
183	cagcagttgc	tattttctca	ggccgggcct	aataacatgt	cggctcaggc	caaaaactgg	1440		
184	ctacccgggc	cctgctaccg	gcagcaacgc	gtctccacga	cactgtcgca	aaataacaac	1500		
185	agcaactttg	cttggaccgg	tgccaccaag	tatcatctga	atggcagaga	ctctctggta	1560		
186	aatcccggtg	tcgctatggc	aacgcacaag	gacgacgaag	agcgatttt	tccatccagc	1620		
187	ggagtcttga	tgtttgggaa	acagggagct	ggaaaagaca	acgtggacta	tagcagcgtt	1680		
188	atgctaacca	gtgaggaaga	aatcaaaacc	accaacccag	tggccacaga	acagtacggc	1740		
189	gtggtggccg						1800		
190	caaggagcct						1860		
191	tgggccaaga						1920		
192	ggactgaaac						1980		
193	ccaactacct						2040		
194	gtcagcgtgg	aaattgaatg	ggagctgcag	aaagagaaca	gcaagcgctg	gaacccagag	2100		

RAW SEQUENCE LISTING

DATE: 08/23/2006 TIME: 10:23:13

Input Set : N:\Crf4\08182006\J573600.raw
Output Set: N:\CRF4\08232006\J573600A.raw

PATENT APPLICATION: US/10/573,600A

195		attcagtata	cttccaacta	taacaaatct	gttaatgtgg	actttactgt	ggacactaat	2160		
196					accagatacc			2217		
198	<210>	SEQ ID NO:								
199	<211>	LENGTH: 221	L7							
200	<212>	TYPE: DNA								
201	<213>	ORGANISM: Unknown								
202	<220>	FEATURE:								
203	<223>	OTHER INFORMATION: new AAV serotype, clone hu.6								
204	<400>	SEQUENCE: 5								
205		atggctgccg	atggttatct	tccagattgg	ctcgaggaca	acctctctga	gggcattcgc	60		
206					aaacccaaag			120		
207		gacggccggg	gtctggtgct	tcctggctac	aagtacctcg	gacccttcaa	cggactcgac	180		
208		aagggggagc	ccgtcaacgc	ggcggacgca	gcggccctcg	agcacgacaa	ggcctacgac	240		
209		cagcagctca	aagcgggtga	caatccgtac	ctgcggtata	accacgccga	cgccgagttt	300		
210		caggagcgtc	tgcaagaaga	tacgtctttt	gggggcaacc	tcgggcgagc	agtcttccag	360		
211		gccaagaagc	gggttctcga	acctctcggt	ctggttgagg	aaggcgctaa	gacggctcct	420		
212		ggaaagaaga	gaccggtaga	gccatcaccc	cagcgttctc	cagactcctc	tacgggcatc	480		
213		ggcaagacag	gccagcagcc	cgcgaaaaag	agactcaact	ttgggcagac	tggcgactca	540		
214		gagtcagtgc	ccgaccctca	accaatcgga	gaaccccccg	caggcccctc	tggtctggga	600		
215		tctggtacaa	tggctgcagg	cggtggcgct	ccaatggcag	acaataacga	aggcgccgac	660		
216		ggagtgggta	gttcctcagg	aaattggcat	tgcgattccg	catggctggg	cgacagagtc	720		
217		atcaccacca	gcacccgacc	ctgggccctc	cccacctaca	acaaccacct	ctacaagcaa	780		
218		atctccaacg	ggacatcggg	aggaagcacc	aacgacaaca	cctacttcgg	ctacagcacc	840		
219		ccctgggggt	attttgactt	taacagattc	cactgccact	tctcaccacg	tgactggcag	900		
220		cgactcatca	acaacaactg	gggattccgg	cccaagagac	tcaacttcaa	gctcttcaac	960		
221		atccaggtca	aggaggtcac	gcagaatgaa	ggcaccaaga	ccatcgccaa	taaccttacc	1020		
222		agcacgattc	aggtctttac	ggactcggaa	taccagctcc	cgtacgtcct	cggctctgcg	1080		
223		caccagggct	gcccgcctcc	gttcccggcg	gacgtcttca	tgattcctca	gtacgggtac	1140		
224		ctgactctga	acaacggcag	tcaggccgtg	ggccgttcct	ccttctactg	cctggagtac	1200		
225		tttccttctc	aaatgcggag	aacgggcaac	aactttgagt	tcagctacca	gtttgaggac	1260		
226		gtgccttttc	acagcagcta	cgcgcatagc	caaagcctgg	accggctgat	gaaccccctc	1320		
227		atcgaccagt	acctgtacta	cctgtctcgg	actcagtcca	cgggaggtac	cgcaggaact	1380		
228		cagcagttgc	tattttctca	ggccgggcct	aataacatgt	cggctcaggc	caaaaactgg	1440		
229		ctacccgggc	cctgctaccg	gcagcaacgc	gtctccacga	cactgtcgca	aaataacaac	1500		
230					tatcatctga			1560		
231					gacgacgaag			1620		
232		ggagtcttga	tgtttgggaa	acagggagct	ggaaaagaca	acgtggacta	tagcagcgtt	1680		
233					accaacccag			1740		
234					gccgctccta			1800		
235					aaccgggacg	_		1860		
236					tttcatcctt			1920		
237					attaagaata			1980		
238			-		tcgttcatca			2040		
239					aaagagaaca			2100		
240					acaaatgtgg			2160		
241		ggtacttatt		ccccattggc	acccgttacc	tcacccgtaa	cctgtaa	2217		
		SEQ ID NO:								
		LENGTH: 221	.7							
245	<212>	TYPE: DNA								

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/573,600A

DATE: 08/23/2006 TIME: 10:23:14

Input Set : N:\Crf4\08182006\J573600.raw
Output Set: N:\CRF4\08232006\J573600A.raw

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 6

VERIFICATION SUMMARY
PATENT APPLICATION: US/10/573,600A

DATE: 08/23/2006 TIME: 10:23:14